



# results of BLAST

BLASTX 2.2.9 [May-01-2004]

**Reference:**

Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

RID: 1088028727-32641-46334205553.BLASTQ4

**Query=**

(759 letters)

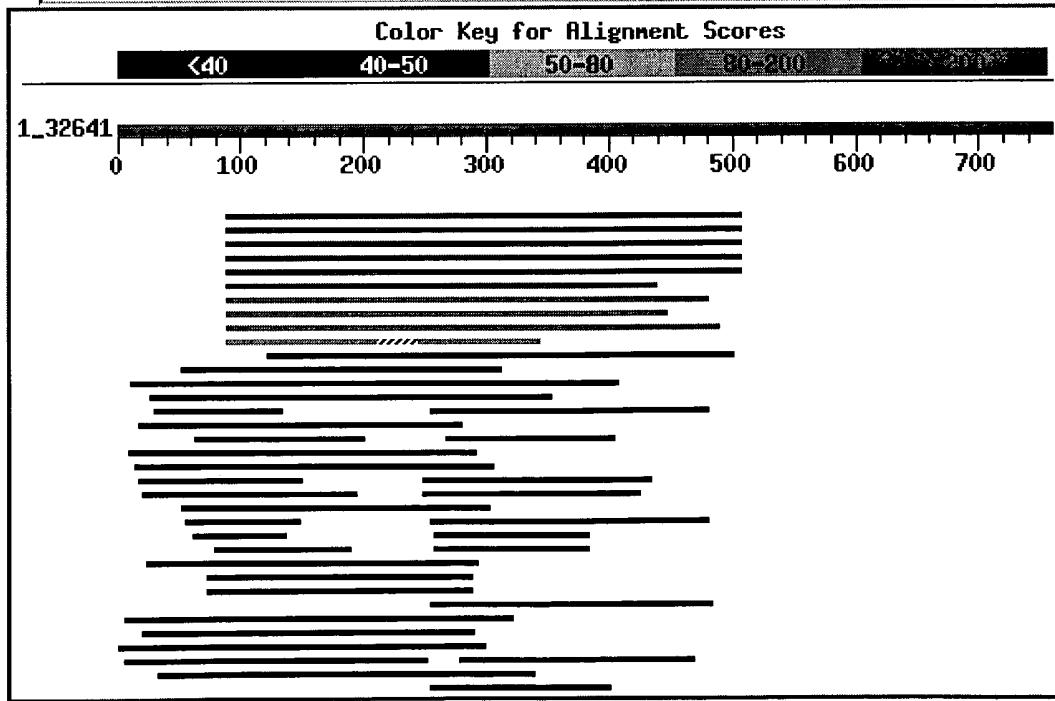
**Database:** All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples  
1,866,121 sequences; 619,474,291 total letters

If you have any problems or questions with the results of this search please refer to the [BLAST FAQs](#)

[Taxonomy reports](#)

## Distribution of 47 Blast Hits on the Query Sequence

Mouse-over to show defline and scores. Click to show alignments



Sequences producing significant alignments:

		Score (bits)	E Value	
gi 34101278 ref NP_898888.1	hypothetical protein MGC10744 ...	280	2e-74	L
gi 34101276 ref NP_115730.2	hypothetical protein MGC10744 ...	272	6e-72	L
gi 21539609 ref NP_080114.1	RIKEN cDNA 1110004B13 [Mus mus...]	257	2e-67	L
gi 27672842 ref XP_213332.1	similar to RIKEN cDNA 1110004B...	257	2e-67	L
gi 47123437 gb AAH70231.1	Unknown (protein for MGC:88213) ...	245	8e-64	
gi 21539635 ref NP_082612.1	RIKEN cDNA 1110004B13 [Mus mus...]	217	2e-55	L
gi 47086013 ref NP_998376.1	zgc:77926 [Danio rerio] >gi 41...	172	8e-42	L
gi 47225380 emb CAG11863.1	unnamed protein product [Tetrao...	169	4e-41	
gi 13751660 gb AAK38512.1	DC20 [Homo sapiens]	119	6e-33	L
gi 34535000 dbj BAC87177.1	unnamed protein product [Homo s...	79	1e-13	
gi 7500843 pir T21990	hypothetical protein F39B2.9 - Caeno...	38	0.25	
gi 46362710 ref ZP_00225559.1	hypothetical protein Krad060...	37	0.33	
gi 18025542 gb AAF78882.2	EBNA-3B [cercopithicine herpesvi...	37	0.56	
gi 41114515 ref XP_371222.1	similar to Hypothetical protei...	36	0.73	L
gi 46321304 ref ZP_00221682.1	COG0477: Permeases of the ma...	36	0.95	
gi 47210644 emb CAG06315.1	unnamed protein product [Tetrao...	35	1.2	
gi 34902106 ref NP_912399.1	unknown protein [Oryza sativa ...]	35	1.2	
gi 7106228 gb AAF36091.1	flagelliform silk protein [Nephil...	35	1.6	
gi 34536289 dbj BAC87601.1	unnamed protein product [Homo s...	35	2.1	
gi 6686017 sp O62732 SYN1_CANFA	Synapsin I >gi 2944066 gb A...	35	2.1	
gi 14701574 dbj BAB62028.1	elastase precursor [Prevotella ...]	35	2.1	
gi 42659679 ref XP_374898.1	similar to pecanex-like 3 [Hom...	34	2.8	L
gi 34365329 emb CAE45990.1	hypothetical protein [Homo sapi...	34	2.8	L
gi 46311822 ref ZP_00212424.1	COG0477: Permeases of the ma...	34	2.8	
gi 13562004 gb AAK30605.1	major ampullate spidroin 2-like ...	34	2.8	
gi 28279307 gb AAH46165.1	Unknown (protein for IMAGE:51920...	34	3.6	
gi 34495966 ref NP_900181.1	probable MFS permease [Chromob...	34	3.6	
gi 37680831 ref NP_935440.1	hypothetical protein VV2647 [V...	34	3.6	
gi 27365118 ref NP_760646.1	Unknown [Vibrio vulnificus CMC...	34	3.6	
gi 24575123 gb AAL06694.1	putative regulatory protein [Str...	34	3.6	
gi 41057143 ref NP_957857.1	ORF080 virion core protein [Or...	34	3.6	
gi 47230172 emb CAG10586.1	unnamed protein product [Tetrao...	33	4.7	
gi 39583515 emb CAE73973.1	Hypothetical protein CBG21601 [...]	33	4.7	
gi 9663052 emb CAC01096.1	110-R orphan receptor [Haemonchu...	33	6.1	
gi 47214200 emb CAG00828.1	unnamed protein product [Tetrao...	33	6.1	
gi 47217474 emb CAG10243.1	unnamed protein product [Tetrao...	33	6.1	
gi 46432590 gb EAK92065.1	hypothetical protein Ca019.6598 ...	33	6.1	
gi 15220207 ref NP_172545.1	expressed protein [Arabidopsis...	33	8.0	
gi 46443901 gb EAL03180.1	hypothetical protein Ca019.4050 ...	33	8.0	
gi 21388708 dbj BAC00817.1	phosphate transport system perm...	33	8.0	
gi 40255174 ref NP_775795.2	chromosome 10 open reading fra...	33	8.0	L
gi 47213146 emb CAF93836.1	unnamed protein product [Tetrao...	33	8.0	
gi 34905662 ref NP_914178.1	P0475H04.14 [Oryza sativa (jap...	33	8.0	

## Alignments

>gi|34101278|ref|NP\_898888.1| L hypothetical protein MGC10744 isoform 2 [Homo sa

gi|37182173|gb|AAQ88889.1| GRVS638 [Homo sapiens]  
Length = 140

Score = 280 bits (716), Expect = 2e-74  
Identities = 140/140 (100%), Positives = 140/140 (100%)  
Frame = +2

Query: 89 MGRVSGLVPSRFLTLLAHLVVVITLFWRSRDSNIQACPLTFTPEEYDKQDIQLVAALSVT 268  
MGRVSGLVPSRFLTLLAHLVVVITLFWRSRDSNIQACPLTFTPEEYDKQDIQLVAALSVT

Sbjct: 1 MGRVSGLVPSRFLTLLAHLVVVITLFWRSRDSNIQACPLTFTPEEYDKQDIQLVAALSVT 60

Query: 269 LGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSAL 448  
LGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSAL

Sbjct: 61 LGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSAL 120

Query: 449 PAVTEMALFVTVFGLKKKPF 508

PAVTEMALFVTVFGLKKKPF

Sbjct: 121 PAVTEMALFVTVFGLKKKPF 140

>gi|34101276|ref|NP\_115730.2|  hypothetical protein MGC10744 isoform 1 [Homo sa  
Length = 146

Score = 272 bits (695), Expect = 6e-72  
Identities = 139/146 (95%), Positives = 140/146 (95%), Gaps = 6/146 (4%)  
Frame = +2

Query: 89 MGRVSGLVPSRFLTLLAHLVVVITLFWRSRDSNIQACPLTFTPEEYDKQDI-----QLV 250  
MGRVSGLVPSRFLTLLAHLVVVITLFWRSRDSNIQACPLTFTPEEYDKQDI +LV

Sbjct: 1 MGRVSGLVPSRFLTLLAHLVVVITLFWRSRDSNIQACPLTFTPEEYDKQDIHPLPLCRLV 60

Query: 251 AALSVTLGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIF 430  
AALSVTLGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIF

Sbjct: 61 AALSVTLGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIF 120

Query: 431 VFCSALPAVTEMALFVTVFGLKKKPF 508

VFCSALPAVTEMALFVTVFGLKKKPF

Sbjct: 121 VFCSALPAVTEMALFVTVFGLKKKPF 146

>gi|21539609|ref|NP\_080114.1|  RIKEN cDNA 1110004B13 [Mus musculus]

gi|12834053|dbj|BAB22768.1|  unnamed protein product [Mus musculus]  
Length = 140

Score = 257 bits (657), Expect = 2e-67  
Identities = 125/140 (89%), Positives = 132/140 (94%)  
Frame = +2

Query: 89 MGRVSGLVPSRFLTLLAHLVVVITLFWRSRDSNIQACPLTFTPEEYDKQDIQLVAALSVT 268  
MGR+SGLVPSRFLTLLAHLVVVITLFWSR+SNIQACPLPL FTPEEY+KQD QLVAAL +T

Sbjct: 1 MGRISGLVPSRFLTLLAHLVVVITLFWRSRESNIQACPLLKFTPEEYEKQDNQLVAALCLT 60

Query: 269 LGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSAL 448  
LGLFAVELAGFLSGVSMFNSTQSL+SI AHCSASVALSFF+FERWECTTYWYIF FCSA

Sbjct: 61 LGLFAVELAGFLSGVSMFNSTQSLLSIAAHCSASVALSFFVFERWECTTYWYIFTFCSAF 120

Query: 449 PAVTEMALFVTVFGLKKKPF 508

PAVTE ALF+ VFGLKKKPF  
 Sbjct: 121 PAVTETALFIAVFGLKKKPF 140

□ >gi|27672842|ref|XP\_213332.1| L similar to RIKEN cDNA 1110004B13 [Rattus norvegicus]  
 Length = 140

Score = 257 bits (657), Expect = 2e-67  
 Identities = 125/140 (89%), Positives = 132/140 (94%)  
 Frame = +2

Query: 89 MGRVSGLVPNSRFLTLLAHLVVVITLFWRSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVT 268  
 MGR+SGLVPSRFLTLLAHLV+VITLFWSR+SNIQACLPL FTPEEY+KQD QLVAAL +T

Sbjct: 1 MGRISGLVPSRFLTLLAHLVIVITLFWRESNIQACLPLKFTPEEYEKQDNQLVAALCLT 60

Query: 269 LGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSAL 448  
 LGLFAVELAGFLSGVSMFNSTQSL+SI AHCSASVALSFFIFERWECTTYWYIF FCSA

Sbjct: 61 LGLFAVELAGFLSGVSMFNSTQSLLSIAAHCSASVALSFFIFERWECTTYWYIFAFCSAF 120

Query: 449 PAVTEMALFVTVFGLKKKPF 508

PAVTE ALF+ VFGLKKKPF

Sbjct: 121 PAVTETALFIAVFGLKKKPF 140

□ >gi|47123437|gb|AAH70231.1| Unknown (protein for MGC:88213) [Homo sapiens]  
 Length = 139

Score = 245 bits (625), Expect = 8e-64  
 Identities = 131/146 (89%), Positives = 132/146 (90%), Gaps = 6/146 (4%)  
 Frame = +2

Query: 89 MGRVSGLVPNSRFLTLLAHLVVVITLFWRSRDSNIQACLPLTFTPEEYDKQDI-----QLV 250  
 MGRVSGLVPNSRFLTLLAHLVVVITLFWRSRDSNIQACLPLTFTPEEYDKQDI +LV

Sbjct: 1 MGRVSGLVPNSRFLTLLAHLVVVITLFWRSRDSNIQACLPLTFTPEEYDKQDIHPLPLCRLV 60

Query: 251 AALSVTLGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIF 430

AALSVTLGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTY

Sbjct: 61 AALSVTLGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTY---- 116

Query: 431 VFCSALPAVTEMALFVTVFGLKKKPF 508

ALPAVTEMALFVTVFGLKKKPF

Sbjct: 117 ---CALPAVTEMALFVTVFGLKKKPF 139

□ >gi|21539635|ref|NP\_082612.1| L RIKEN cDNA 1110004B13 [Mus musculus]

gi|12850002|gb|BAB28560.1| L unnamed protein product [Mus musculus]

Length = 143

Score = 217 bits (552), Expect = 2e-55  
 Identities = 105/117 (89%), Positives = 111/117 (94%)  
 Frame = +2

Query: 89 MGRVSGLVPNSRFLTLLAHLVVVITLFWRSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVT 268  
 MGR+SGLVPSRFLTLLAHLVVVITLFWSR+SNIQACLPL FTPEEY+KQD QLVAAL +T

Sbjct: 1 MGRISGLVPSRFLTLLAHLVVVITLFWRESNIQACLPLKFTPEEYEKQDNQLVAALCLT 60

Query: 269 LGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFC 439  
 LGLFAVELAGFLS VSMFNSTQSL+SI AHCSASVALSFF+FERWECTTYWYIF FC  
 Sbjct: 61 LGLFAVELAGFLSRVSMFNSTQSLLSIAAHCSASVALSFFVFERWECTTYWYIFTFC 117

□ >gi|47086013|ref|NP\_998376.1| L zgc:77926 [Danio rerio]  
 gi|41351480|gb|AAH65897.1| L Zgc:77926 [Danio rerio]  
 Length = 135

Score = 172 bits (435), Expect = 8e-42  
 Identities = 80/131 (61%), Positives = 100/131 (76%)  
 Frame = +2

Query: 89 MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVT 268  
 M + LVP+RFLTL AHLV++IT+FWSRD+NIQ+CLPL FT ++Y +D +L ALSVT  
 Sbjct: 1 MSALKSLVPARFLTLTAHLVIIITIFWSRDNNIQSCLPLEFTEDQYRTEDETRLTVALSVT 60

Query: 269 LGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSAL 448  
 L LF +ELAGFLSGVSMFN Q+L+S+ H SA V LSFF+F +W C TYW IF CS  
 Sbjct: 61 LALFVLELAGFLSGVSMFNSQALLSLITHSSACVCLSFFVHQWPCWTYWIIFSICSVF 120

Query: 449 PAVTEMALFVT 481  
 PAV E+ L ++  
 Sbjct: 121 PAVVELFLLS 131

□ >gi|47225380|emb|CAG11863.1| unnamed protein product [Tetraodon nigroviridis]  
 Length = 180

Score = 169 bits (429), Expect = 4e-41  
 Identities = 77/120 (64%), Positives = 100/120 (83%)  
 Frame = +2

Query: 89 MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVT 268  
 M VS LVP+RFLT++AHLV+VIT+FWSR++N++A LPL FT E+YD +D +LV AL+VT  
 Sbjct: 1 MPAVSSLVPARFLTIIAHLVIVITIFWSRENNVRAGLPLDFTQEYDSEDRKLVIALAVT 60

Query: 269 LGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSAL 448  
 +G+FA+ELAGF SGVSMFN +Q L+S G H SASVAL FF+FE+WEC YW+IFV C ++  
 Sbjct: 61 IGMFAIELAGFFSGVSMFNCSQGLLSTGTHASASVALFFLFEQWECDIYWWIFVICRSV 120

□ >gi|13751660|gb|AAK38512.1| L DC20 [Homo sapiens]  
 Length = 141

Score = 119 bits (297), Expect = 9e-26  
 Identities = 70/113 (61%), Positives = 74/113 (65%), Gaps = 6/113 (5%)  
 Frame = +2

Query: 89 MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDI-----QLV 250  
 MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDI +LV  
 Sbjct: 1 MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIHPLPLCRLV 60

Query: 251 AALSVTILGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWEC 409  
 AALSVTILG SG F + + H AS L + W C  
 Sbjct: 61 AALSVTILG-----PLCSGAGRFPLRS--LHVQQHPRASSPLGLTVVHPWPC 104

Score = 114 bits (284), Expect(2) = 6e-33  
Identities = 54/54 (100%), Positives = 54/54 (100%)  
Frame = +3

Query: 330 PRASSPLGLTVVHPWPCPSSYSSVGSALRIGTFLSSAVPFQLSLKWLWLYSSPSLG 491  
PRASSPLGLTVVHPWPCPSSYSSVGSALRIGTFLSSAVPFQLSLKWLWLYSSPSLG  
Sbjct: 88 PRASSPLGLTVVHPWPCPSSYSSVGSALRIGTFLSSAVPFQLSLKWLWLYSSPSLG 141

Score = 50.1 bits (118), Expect(2) = 6e-33  
Identities = 21/21 (100%), Positives = 21/21 (100%)  
Frame = +1

Query: 271 GPLCSGAGRFPLRSLHVQQHP 333  
GPLCSGAGRFPLRSLHVQQHP  
Sbjct: 68 GPLCSGAGRFPLRSLHVQQHP 88

>gi|34535000|dbj|BAC87177.1| unnamed protein product [Homo sapiens]  
Length = 163

Score = 79.0 bits (193), Expect = 1e-13  
Identities = 33/33 (100%), Positives = 33/33 (100%)  
Frame = +3

Query: 246 WWPRSLSPWASLQWSWPVSSQESPCSTAPRASS 344  
WWPRSLSPWASLQWSWPVSSQESPCSTAPRASS  
Sbjct: 84 WWPRSLSPWASLQWSWPVSSQESPCSTAPRASS 116

Score = 60.8 bits (146), Expect = 3e-08  
Identities = 32/41 (78%), Positives = 32/41 (78%)  
Frame = +2

Query: 89 MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTF 211  
MGRVSGLVPSRFLTLLAHLVVVITLFWSR A P F  
Sbjct: 1 MGRVSGLVPSRFLTLLAHLVVVITLFWSRVRPTAALNPSPF 41

>gi|7500843|pir|T21990 hypothetical protein F39B2.9 - Caenorhabditis elegans  
Length = 136

Score = 37.7 bits (86), Expect = 0.25  
Identities = 31/127 (24%), Positives = 58/127 (45%)  
Frame = +2

Query: 122 FLTLLAHLVVVITLFWSRDSNIQACLPLTFPEEYDKQDIQLVAALSVTLGLFAVELAGF 301  
F L+ H + +S+ S+++A + + E + I L T +E A  
Sbjct: 12 FQCLMGHFA LSTITIFSQASHLEASVSGGASISEEAR--IGFTVCLVFTKLAILIEFASI 69

Query: 302 LSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFVT 481  
L V + + ++SL S +H AS+ FI++ +W +F F S P V +A+ +

Sbjct: 70 L--VDLPSPSRSLYSTLSHTIASIFFLIFIYDSHPPAHFWLLFAFLSLPPCV--LAVSAS 125

Query: 482 VFGLKKK 502

+ G ++K

Sbjct: 126 IGGFRRK 132

>gi|46362710|ref|ZP\_00225559.1| hypothetical protein Krad06004793 [Kineococcus r  
SRS30216]  
Length = 293

Score = 37.4 bits (85), Expect = 0.33

Identities = 31/87 (35%), Positives = 41/87 (47%)

Frame = +1

Query: 52 PASETAALVPEGHGPGLRACALSLPDAPGASGGRHHLILVPGQQHTGLPASHVHPRGV\*Q 231

P + A+ PEG G GL A+ L + PGA G +H +L GQ+ P H GV

Sbjct: 57 PVPDGVAHVPEGVGGGLPG-AVVLEEGPGA-GPQHGPVLDRGQRRQPPRPRPHRAGVQD 114

Query: 232 AGHSAGGRALCHPGPLCSGAGRFLRS 312

G AG PGP G+ P+R+

Sbjct: 115 LGQRAG-----PGP-----GQRPVRA 130

>gi|18025542|gb|AAF78882.2| EBNA-3B [cercopithicine herpesvirus 15]  
Length = 928

Score = 36.6 bits (83), Expect = 0.56

Identities = 42/149 (28%), Positives = 50/149 (33%), Gaps = 17/149 (11%)

Frame = -2

Query: 407 TPNARI\*RRTGPR-----MHYSEPQWR\*GSGCC\*TWRLLRGNRPAPLQRGPG\*QR---- 258

TP A I + PR + + P W + N P+ PG R

Sbjct: 601 TPRAHIPANSDPRRATFDVLGFRSPDWPPKNWTNANPSTANANASRPISAAPGRARPPAA 660

Query: 257 ----ARPPAECPACHTPRG\*T\*EAGRPVCCCPGTRIR\*\*RPPDAPGASGSERAQALRPG- 93

ARPPA P P RP PG RPP A A G R A PG

Sbjct: 661 APGRARPPAAAPGRARPPAAATGRARPPAAAPGRA---RPPAA--APGRARPPAAAPGR 714

Query: 92 --PWPSGTRA AVSEAGSSETATPEVSPQA 12

P + A A + A P+ PQA

Sbjct: 715 ARPPAAAPGRARPPAAAPGRAAPQQQPQA 743

>gi|41114515|ref|XP\_371222.1| L similar to Hypothetical protein MGC37938 [Homo s  
Length = 215

Score = 36.2 bits (82), Expect = 0.73

Identities = 34/110 (30%), Positives = 45/110 (40%), Gaps = 1/110 (0%)

Frame = +1

Query: 28 SGVAVSELPASETAALVPEGHG-PGLRACALSLPDAPGASGGRHHLILVPGQQHTGLPAS 204

+ + V E + L+ G G PG R A +LP P +G R L P G P+

Sbjct: 92 AALVVLEFSLRVSTLLSLGKGLPGCRREAAALPAVPVLAGLRADLW--PELPAGGRPSP 149

Query: 205 HVHPRGV\*QAGHSAGGRALCHPGPLCSGAGRFLRSIHVQQHPEPHLHWG 354

H P +AGH+AG P P PLR+ +P L WG  
 Sbjct: 150 HAEPAAEPRAGHAAGPGCPAPPPPRLP----PLRAAQ----QPALLWG 189

►>gi|46321304|ref|ZP\_00221682.1| COG0477: Permeases of the major facilitator superfamily [Burkholderia cepacia R1808]  
 Length = 428

Score = 35.8 bits (81), Expect = 0.95  
 Identities = 22/77 (28%), Positives = 38/77 (49%), Gaps = 1/77 (1%)  
 Frame = +2

Query: 254 ALSVTLGLFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFV 433  
 A+SV + L+ + + + F +++++ SA VALS + + W+ W + V  
 Sbjct: 57 AISVNIALYGLTGPFAAAAMQRFGLRPTILTAVALTLSAGVALSSMMTQSWMQMVVWGLMV 116

Query: 434 FCS-ALPAVTEMALFVT 481  
 CS + A+T A FVT  
 Sbjct: 117 GCSTGVVALTLSATFVT 133

►>gi|47210644|emb|CAG06315.1| unnamed protein product [Tetraodon nigroviridis]  
 Length = 325

Score = 35.4 bits (80), Expect = 1.2  
 Identities = 28/88 (31%), Positives = 41/88 (46%)  
 Frame = -3

Query: 280 KEAQGDRERGHQLNVLLVILLGGERERQAGLYVAVPGPE\*GDDDHQMRQERQEARGHKP\* 101  
 + A+G R + H+L+ + + GG + GL VPGP DH++R+ER +G  
 Sbjct: 203 RRARGRRGQEHLHAGVPLGQGGLQ----GLRRPVPGPVRALRDHEVREERARGQGRAVR 258

Query: 100 DPAHGPRGQGRRSLRLEVQRQRLLKSPR 17  
 A P G R QR + L PR  
 Sbjct: 259 AGAEPPAGGAHRR----QRAQLHPPR 281

►>gi|34902106|ref|NP\_912399.1| unknown protein [Oryza sativa (japonica cultivar-group)]  
 ►>gi|29893617|gb|AAP06871.1| unknown protein [Oryza sativa (japonica cultivar-group)]  
 Length = 279

Score = 35.4 bits (80), Expect = 1.2  
 Identities = 18/46 (39%), Positives = 25/46 (54%)  
 Frame = +3

Query: 267 PWASLQWSWPVSSQESPCSTAPRASSPLGLTVVHPWPCPSSYSSVG 404  
 PW SL+ S PV +E P + AP + + GLT+ W P +VG  
 Sbjct: 164 PWLSLRNSCPVCRRELPAAAAPESEADAGLTI---WRLPRGGFAVG 206

►>gi|7106228|gb|AAF36091.1| flagelliform silk protein [Nephila madagascariensis]  
 Length = 1884

Score = 35.0 bits (79), Expect = 1.6  
 Identities = 28/96 (29%), Positives = 35/96 (36%), Gaps = 2/96 (2%)

Frame = +1

Query: 10 SACGETSGVAVSELPASETAALVPEGHGPGLRACALSLPDAPGASGGRHHHLILVPGQQHT 189  
 S G+ SG S P G GPG + S P G SG R G +  
 Sbjct: 118 SGAGQPSGAGPSGTGGYAPTYAPSGSGPGVVRPSASGPSSGSGPSGSRPSSSGSSGTRPS 177

Query: 190 GLPASHVHPRGV\*QAGHSAGGRAL--CHPGPLCSGA 291  
 A P G+ G S GG + GP SG+  
 Sbjct: 178 ANAAGGSSPGGIAPGGSSPGGAGVSGATGGPASSGS 213

□ >gi|34536289|dbj|BAC87601.1| unnamed protein product [Homo sapiens]  
 Length = 181

Score = 34.7 bits (78), Expect = 2.1  
 Identities = 26/76 (34%), Positives = 30/76 (39%), Gaps = 14/76 (18%)  
 Frame = +3

Query: 249 WPRSLS-----PWASLQWSWPVSSQESPCSTAPRASSPLGLT---VVHPWPCPS---- 386  
 WP S+ P +L W WP P A S L L V PWPCP+  
 Sbjct: 102 WPCSVFGRALCYPSPALPWYWHLPPWPGPTLAFSTLALPFPGVALPWPCPALALAL 161

Query: 387 SYSSVGSALRIGHTFLS 434  
 Y GSAL + LS  
 Sbjct: 162 PYPGPGSALTALVL 177

□ >gi|6686017|sp|O62732|SYN1\_CANFA Synapsin I  
 gi|2944066|gb|AAC05207.1| synapsin I [Canis familiaris]  
 Length = 415

Score = 34.7 bits (78), Expect = 2.1  
 Identities = 31/99 (31%), Positives = 43/99 (43%), Gaps = 2/99 (2%)  
 Frame = -2

Query: 305 RGNRPAPLQRGPG\*QRARPPAECPACHTPRG\*T\*EAGRPVCCCPGTRIR\*\*RPPDAPGA- 129  
 +G + P+ GPG A P A PA +P+ +AG P T + PP A GA  
 Sbjct: 257 QGRQSRPVGAGPG---APPAARPPASPSPQR---QAGPPQATRQ-TSVSGQAPPKASGAP 309

Query: 128 -SGSERAQALRPGPWPSGTRAAVSEAGSSETATPEVSPQ 15  
 SG +R + P P+G S+AG P + Q  
 Sbjct: 310 PSGQQRQGPPQKPPGPAGPTRQASQAGPMPRTGPPTTQQ 348

□ >gi|14701574|dbj|BAB62028.1| elastase precursor [Prevotella intermedia]  
 Length = 640

Score = 34.7 bits (78), Expect = 2.1  
 Identities = 18/59 (30%), Positives = 25/59 (42%)  
 Frame = +3

Query: 249 WPRSLSPWASLQWSWPVSSQESPCSTAPRASSPLGLTVVHPWPCPSSYSSVGSALRIGHT 425  
 WP + PW WP + P P P+G + P+P P Y +G +R GT  
 Sbjct: 265 WPWPVDPW-----WPPIPKPWP----PIGPGPIGPYPPPRPYRPIGPRIRFGT 313

□ >gi|42659679|ref|XP\_374898.1| L similar to pecanex-like 3 [Homo sapiens]  
Length = 2064

Score = 34.3 bits (77), Expect = 2.8  
Identities = 19/44 (43%), Positives = 27/44 (61%), Gaps = 2/44 (4%)  
Frame = +3

Query: 258 SLSPWASLQWSWPVSSQESPC--STAPRASSPLGLTVVHPWPCP 383  
SLSP S + S P +SQ+ PC S+AP + +P+G + WP P  
Sbjct: 2006 SLSPDVSTEASPPRASQDIPCLDSSAPESGTPMG--ALGDPAP 2047

□ >gi|34365329|emb|CAE45990.1| L hypothetical protein [Homo sapiens]  
Length = 924

Score = 34.3 bits (77), Expect = 2.8  
Identities = 19/44 (43%), Positives = 27/44 (61%), Gaps = 2/44 (4%)  
Frame = +3

Query: 258 SLSPWASLQWSWPVSSQESPC--STAPRASSPLGLTVVHPWPCP 383  
SLSP S + S P +SQ+ PC S+AP + +P+G + WP P  
Sbjct: 866 SLSPDVSTEASPPRASQDIPCLDSSAPESGTPMG--ALGDPAP 907

□ >gi|46311822|ref|ZP\_00212424.1| COG0477: Permeases of the major facilitator superfamily [Burkholderia cepacia R18194]  
Length = 427

Score = 34.3 bits (77), Expect = 2.8  
Identities = 21/77 (27%), Positives = 38/77 (49%), Gaps = 1/77 (1%)  
Frame = +2

Query: 254 ALSVTLGLFAVELAGFLSGVSMFNSTQLISIGAHCASVALSFFIFERWECTTYWYIFV 433  
A+SV + L+ + + + F +++++ A VALS + + W+ W + V  
Sbjct: 57 AISVNIALYGLTGPFAAAAMQRGLRPTILTAALVTMCAGVALSSMMTQSWQMVVIWGLMV 116

Query: 434 FCSA-LPAVTEMALFVT 481  
CS+ + A+T A FVT  
Sbjct: 117 GCSSGVVVALTLSATFVT 133

□ >gi|13562004|gb|AAK30605.1| major ampullate spidroin 2-like protein [Nephila madagascariensis]  
Length = 1953

Score = 34.3 bits (77), Expect = 2.8  
Identities = 26/85 (30%), Positives = 36/85 (42%), Gaps = 1/85 (1%)  
Frame = +1

Query: 52 PASETAALVPEGHGPGLRACALSLPDAPGASGRRHHLILVPGQQHTGLPASHVHPRGV\*Q 231  
P + AA P G+GPG + + A G+ G + PGQQ G P + G  
Sbjct: 1562 PGAAAAAAAGPGGYGPGQQPGAAAAAAAGSGPGGGYG---PGQQGPGPGAAAAAAGRGP 1617

Query: 232 AGHSAGGRALCHPG-PLCSGAGRFP 303  
G+ G + PG + AGR P  
Sbjct: 1618 GGYGPGQQGPGGPGAAAAAAAGRGP 1642

Score = 34.3 bits (77), Expect = 2.8  
 Identities = 26/85 (30%), Positives = 36/85 (42%), Gaps = 1/85 (1%)  
 Frame = +1

Query: 52 PASETAALVPEGHGPGLRACALSLPDAPGASGGRHHLILVPGQQHTGLPASHVHPRGV\*Q 231  
 P + AA P G+GPG + + A G+ G + PGQQ G P + G  
 Sbjct: 951 PGAAAAAAAGPGGGYGPQQGPQGAAAAAAAGSGPGGGYG---PGQQGPGGPGAAAAAAAGRGP 1006

Query: 232 AGHSAGGRALCHPG-PLCSGAGRFP 303  
 G+ G + PG + AGR P  
 Sbjct: 1007 GGYGPGQQGPGGPGAAAAAAAGRGP 1031

□>gi|28279307|gb|AAH46165.1| Unknown (protein for IMAGE:5192014) [Homo sapiens]  
 □>gi|29387251|gb|AAH48278.1| Unknown (protein for IMAGE:5192302) [Homo sapiens]  
 Length = 205

Score = 33.9 bits (76), Expect = 3.6  
 Identities = 20/40 (50%), Positives = 21/40 (52%), Gaps = 5/40 (12%)  
 Frame = -2

Query: 134 GASGSERAQAL-RPGPWPSGTRAAVSEAGS----SETATP 30  
 GA G ER L RPG WP G RA +A S S T TP  
 Sbjct: 59 GAGGGEREDGLWRPGRWPRGARAPGEQAASRSPWSATGTP 98

□>gi|34495966|ref|NP\_900181.1| probable MFS permease [Chromobacterium violaceum A  
 □>gi|34101820|gb|AAQ58188.1| probable MFS permease [Chromobacterium violaceum ATCC  
 Length = 425

Score = 33.9 bits (76), Expect = 3.6  
 Identities = 23/77 (29%), Positives = 36/77 (46%)  
 Frame = +2

Query: 254 ALSVTLGLFAVELAGFLSGVSMFNSTQLISIGAHCSASVALSFFIFERWECTTYWYIFV 433  
 ALS+ L LF + + + F +++++S A ASVALS + W+ W + V  
 Sbjct: 57 ALSINLALFGLMGPAAAAMLRFGLRRTVLSALALLGASVALSTLMRSNWQLQLLWGVMV 116

Query: 434 FCSALPAVTEMALFVTV 484  
 C+ T M L +V  
 Sbjct: 117 GCAT--GATAMTLGASV 131

□>gi|37680831|ref|NP\_935440.1| hypothetical protein VV2647 [Vibrio vulnificus YJ0  
 □>gi|37199580|gb|BAC95411.1| hypothetical protein [Vibrio vulnificus YJ016]  
 Length = 464

Score = 33.9 bits (76), Expect = 3.6  
 Identities = 29/75 (38%), Positives = 39/75 (52%), Gaps = 3/75 (4%)  
 Frame = +2

Query: 74 LSPRAMGRVSGLVPSRFLTLAHLVVVITLFWSRDSNIQACLPLTFT-PEEYDKQDIQLV 250  
 LS A VSG++ S L+A VV LFWSR S A LPL P + + L+

Subjct: 361 LSGIAYSLVSGVLDSQLMA--VVACALFWSRCASASLPLPANKPRRWWAHGVILI 418

Query: 251 AAL--SVTLGLFAVE 289

A+ SV++G A+E

Subjct: 419 IAIGCSVSVGYRALE 433

□ >gi|27365118|ref|NP\_760646.1| Unknown [Vibrio vulnificus CMCP6]

gi|27361264|gb|AAO10173.1| Unknown [Vibrio vulnificus CMCP6]

Length = 464

Score = 33.9 bits (76), Expect = 3.6

Identities = 29/75 (38%), Positives = 39/75 (52%), Gaps = 3/75 (4%)

Frame = +2

Query: 74 LSPRAMGRVSGLVPSRFLTLAHLVVVITLFWSRDSNIQACLPLTFT-PEEYDKQDIQLV 250

LS A VSG++ S L+A VV LFWSR S A LPL P + + L+

Subjct: 361 LSGIAYSLVSGVLDSQLMA--VVACALFWSRCASASLPLPANKPRRWWAHGVILI 418

Query: 251 AAL--SVTLGLFAVE 289

A+ SV++G A+E

Subjct: 419 IAIGCSVSVGYRALE 433

□ >gi|24575123|gb|AAL06694.1| putative regulatory protein [Streptomyces globisporu

Length = 196

Score = 33.9 bits (76), Expect = 3.6

Identities = 23/52 (44%), Positives = 26/52 (50%), Gaps = 6/52 (11%)

Frame = +1

Query: 64 TAALVPEGHGPGLRACALSLPDAPGA-----SGGRHHLILVPGQQHTGLPA 201

TA + +GH PGL A LS+PDA G S GR L P TG PA

Subjct: 146 TAQAMEDGH-PGLEAMGLSMPDADGLWRFKDYSYGREFARLHPDTTSTGSPA 196

□ >gi|41057143|ref|NP\_957857.1| ORF080 virion core protein [Orf virus]

gi|41018700|gb|AAR98305.1| ORF080 virion core protein [Orf virus]

Length = 328

Score = 33.9 bits (76), Expect = 3.6

Identities = 28/94 (29%), Positives = 34/94 (36%), Gaps = 4/94 (4%)

Frame = -2

Query: 293 PAPLQRGPG\*QRARPPAECPACHTPRG\*T\*EAGRPVCCCPGTRIR\*\*RPP---DAPGAS 126

PAP P A PPA P C P A C P TR PP AP A

Subjct: 155 PAPAVTCPAPAPACPPATAPTCPPAVCPAPARPPAPACPPSTRQCPPAPPLPTKPAPAAK 214

Query: 125 GSERAQALRPGPWPSGTRAAVSEAGSSETATPEV 24

Q+L P +P+ + + +S P V

Subjct: 215 PIFLQQSLPPPQYPASSCPTIKAPAASPVLEPRV 248

□ >gi|47230172|emb|CAG10586.1| unnamed protein product [Tetraodon nigroviridis]

Length = 441

Score = 33.5 bits (75), Expect = 4.7  
 Identities = 40/118 (33%), Positives = 48/118 (40%), Gaps = 13/118 (11%)  
 Frame = +1

Query: 7 LSACGETSGVAVSELPASETAALVPEGH--GPGLRACALSLPDAPGA--SGGRHHHLILVP 174  
 L A GE A +P H GP RA P+APG+ +G + VP  
 Sbjct: 256 LPAGGEREQRAAGRVPGGPLRGHPAVAHQRGPLPRAH---PEAPGSPQAGDQRGAAPVP 311

Query: 175 -GQQHTGLPASHVHPRGV\*QAGHSAGGRALCHP-----GPLCSGAGRFPRLRSLHV 321  
 G QH GLP PR + AG +A G P PL +GA R P R HV  
 Sbjct: 312 RGAQHAGLPQHAAQPRPLLPAGQAALGLPGLRPRARLSQLGQPLGAGAQR-PARVPHV 368

>gi|39583515|emb|CAE73973.1| Hypothetical protein CBG21601 [Caenorhabditis briggsae]  
 Length = 719

Score = 33.5 bits (75), Expect = 4.7  
 Identities = 24/90 (26%), Positives = 30/90 (33%)  
 Frame = -2

Query: 290 APLQRGPG\*QRARPPAECPACHTPRG\*T\*EAGRPVCCCPGTRIR\*\*RPPDAPGASGSERA 111  
 A R P + + P A P PR A P PG AP AS +  
 Sbjct: 548 ASAPRAPAPRASAPRAPAPRASAPRASAPRALAPGA-----SAPSASAPRAS 599

Query: 110 QALRPGPWPSGTRAAVSEAGSSETATPEVS 21  
 P S TRA+ A + + P S  
 Sbjct: 600 APRASAPRASATRASAPRASAPRASAPRAS 629

>gi|9663052|emb|CAC01096.1| 110-R orphan receptor [Haemonchus contortus]  
 Length = 986

Score = 33.1 bits (74), Expect = 6.1  
 Identities = 16/44 (36%), Positives = 23/44 (52%)  
 Frame = -2

Query: 149 PPDAPGASGSERAQALRPGPWPSGTRAAVSEAGSSETATPEVSP 18  
 PP P A G+ ++ +RP +A S+ GSS T EV+P  
 Sbjct: 921 PPPPPPAQGTTPSKVKRPPSSKMSDDSSAYSDGGSSSVLTTEVTP 964

>gi|47214200|emb|CAG00828.1| unnamed protein product [Tetraodon nigroviridis]  
 Length = 2135

Score = 33.1 bits (74), Expect = 6.1  
 Identities = 34/105 (32%), Positives = 42/105 (40%), Gaps = 5/105 (4%)  
 Frame = +1

Query: 1 LDLSACGETSGVAVSELPASETAALVPEGH--GPGLRACALSLPDAPGASGGRHHHLILVP 174  
 L L + G +SG + LP PE H GP L LS P PG S L VP  
 Sbjct: 1484 LGLPSAGASSGSSAFLPP-----PELHLRGPSLSQNKLSPHEPGLSDSPSSLNPV 1536

Query: 175 GQ-QHTGLPASHVHPRGV\*QAGHSAGGRALCHP--GPLCSGAGRFPRLRSLHV 300  
 H + +SH+ P G + H A HP P G+ F  
 Sbjct: 1537 SPGTHFHISSHLQPSG---SSHPOQLSLAGLHPLLPPAGGGSAEF 1578

►>gi|47217474|emb|CAG10243.1| unnamed protein product [Tetraodon nigroviridis]  
Length = 1190

Score = 33.1 bits (74), Expect = 6.1  
Identities = 22/62 (35%), Positives = 28/62 (45%), Gaps = 4/62 (6%)  
Frame = -2

Query: 194 RPVCCCPGTRIR\*\*RPPDAPGA----SGSERAQALRPGPWPSGTRAAVSEAGSSETATPE 27  
RP CP R RPP APGA GS A GP +G + GS+E+ P  
Sbjct: 97 RPRASCPRLLR----RPPGAPGAHCSRQGSAEASPSWAGPSHAGFLPPTRKPGSNESIPPA 152

Query: 26 VS 21  
++  
Sbjct: 153 IT 154

►>gi|46432590|gb|EAK92065.1| hypothetical protein Ca019.6598 [Candida albicans SC]  
Length = 781

Score = 33.1 bits (74), Expect = 6.1  
Identities = 28/85 (32%), Positives = 34/85 (40%), Gaps = 3/85 (3%)  
Frame = -2

Query: 251 PPAEC-PACHTPRG\*T\*EAGRPVCCCPGT--RIR\*\*RPPDAPGASGSERAQALRPGPWPS 81  
PPA P PR T + G P PG R +PP AP S A P P P  
Sbjct: 475 PPARATPPAPPRARTNQLGLPPRNTGLPPRTNNTQPPPAPPRASRGAVPPPPPPPR 534

Query: 80 GTRAADVSEAGSSETATPEVSPQADK 6  
TRA + S + +SP A +  
Sbjct: 535 ATRAPMQLQLQSSPQSSPISPPAQ 559

►>gi|15220207|ref|NP\_172545.1| expressed protein [Arabidopsis thaliana]  
gi|28393599|gb|AAO42219.1| putative carboxyl-terminal peptidase [Arabidopsis thal  
gi|29824323|gb|AAP04122.1| putative carboxyl-terminal peptidase [Arabidopsis thal  
Length = 467

Score = 32.7 bits (73), Expect = 8.0  
Identities = 15/31 (48%), Positives = 16/31 (51%)  
Frame = -3

Query: 148 HQMRQERQEARGHKP\*DPAHGPRGQGRRSLR 56  
HQ + RG KP DP PRG RR LR  
Sbjct: 135 HQPAFDHPSLRGQKPLDPPERPRGHNRRGLR 165

►>gi|46443901|gb|EAL03180.1| hypothetical protein Ca019.4050 [Candida albicans SC]  
gi|46444066|gb|EAL03344.1| hypothetical protein Ca019.11532 [Candida albicans SC5  
Length = 154

Score = 32.7 bits (73), Expect = 8.0  
Identities = 16/49 (32%), Positives = 25/49 (51%)  
Frame = +3

Query: 255 RSLSPWASLQWSWPVSSQESPCSTAPRASSPLGLTVVHPWPCPSSYSSV 401  
 R +S + +W +S S + +P SSP T++ WPC YSS+  
 Sbjct: 2 REMSSMSESSLAWTAASSSLVTNSPSFWPNTLIRNWPC--GYSSL 48

□>gi|21388708|dbj|BAC00817.1| phosphate transport system permease ABC transporter  
 [Rhodobacter sphaeroides f. sp. denitrificans]  
 Length = 481

Score = 32.7 bits (73), Expect = 8.0  
 Identities = 31/102 (30%), Positives = 40/102 (39%)  
 Frame = +1

Query: 34 VAVSELPASETAALVPEGHGPGLRACALSLPDAPGASGGRHHLILVPGQQHTGLPASHVH 213  
 V L + AAL PEG P LRA ++ D + L + H L AS  
 Sbjct: 105 VVAEGLSDRDLAALTPEGLAPRLRAVGAADVADITPQT-----LDAARAHRALQASFGL 157

Query: 214 PRGV\*QAGHSAGGRALCHPGPLCSGAGRFLRSLHVQQHPEP 339  
 V G S GGR+ PGP AG + + H +P  
 Sbjct: 158 WMSVVLVGLGASLGGRS---PGPPAHSAGLPCPQRVRALHHVDP 196

□>gi|40255174|ref|NP\_775795.2| L chromosome 10 open reading frame 64 [Homo sapien  
gi|34534655|dbj|BAC87073.1| L unnamed protein product [Homo sapiens]  
 Length = 1042

Score = 32.7 bits (73), Expect = 8.0  
 Identities = 13/25 (52%), Positives = 16/25 (64%)  
 Frame = -2

Query: 137 PGASGSERAQALRPGPWPSGTRAAV 63  
 PG SGS+ AQ L GPWP+ A +  
 Sbjct: 951 PGCSSQTAQGLAEGPWPAAAPDAGL 975

□>gi|47213146|emb|CAF93836.1| unnamed protein product [Tetraodon nigroviridis]  
 Length = 708

Score = 32.7 bits (73), Expect = 8.0  
 Identities = 18/38 (47%), Positives = 20/38 (52%), Gaps = 1/38 (2%)  
 Frame = +1

Query: 79 PEGHGPGLRACALSLPDAPGA-SGGRHHLILVPGQQHT 189  
 PEG GP +A A S P+ P GGR L P Q HT  
 Sbjct: 508 PESEGPRTKAGACSTPEGPATPPGGRKILKKTPEQLHT 545

□>gi|34905662|ref|NP\_914178.1| P0475H04.14 [Oryza sativa (japonica cultivar-group)  
gi|12060519|dbj|BAB20648.1| P0475H04.14 [Oryza sativa (japonica cultivar-group)]  
 Length = 529

Score = 32.7 bits (73), Expect = 8.0  
 Identities = 22/66 (33%), Positives = 27/66 (40%), Gaps = 2/66 (3%)

Frame = +3

Query: 279 LQWSWPVSS--QESPCSTAPRASSPLGLTVVHPWPCPSSYSSVGSALRIGTFLSSAVPFQ 452  
L WS+P S + PC P+ P PCP Y V AL + SA P  
Sbjct: 68 LSWSFPPSIVLEHHPCENFSFFPPPIDRKRTGPRPCPVCYVPVEQALALMPGAPSASPVL 127

Query: 453 LSLKWL 470

SL +L

Sbjct: 128 RSLNYL 133

Lambda K H  
0.318 0.134 0.401

Gapped

Lambda K H  
0.267 0.0410 0.140

Matrix: BLOSUM62

Gap Penalties: Existence: 11, Extension: 1

Number of Sequences: 1866121

Number of Hits to DB: 192,225,388

Number of extensions: 5047647

Number of successful extensions: 16879

Number of sequences better than 10.0: 9

Number of HSP's better than 10.0 without gapping: 15304

Number of HSP's gapped: 16825

Number of HSP's successfully gapped: 10

Number of extra gapped extensions for HSPs above 10.0: 0

Length of query: 759

Length of database: 619,474,291

Length adjustment: 122

Effective length of query: 637

Effective length of database: 391,807,529

Effective search space: 249581395973

Effective search space used: 51326786299

T: 12

A: 40

X1: 16 ( 7.3 bits)

X2: 38 (15.0 bits)

X3: 64 (25.0 bits)

S1: 41 (22.0 bits)

S2: 72 (32.3 bits)